Watcha say?- developing and implementing standards for assessing speaking

Dr. Craig L. Dicker

Office of English Language Programs

US Embassy, Ankara

DickerCL@State.gov

http://turkey.embassy.gov/public_affairs.html

Competence vs. Performance

- Definition of "Competence"- What a person can actually do
- Can NOT be directly measured
- A purely psychological construct- no reference to the social nor environmental variables

- Definition of "Performance"- what a person actually does in a particular setting (e.g. taking a test)
- Competence can only be measured via one's performances
- One's performance can never equal one's ideal competence- the former will always fall somewhat short

Responsibilities of Test Givers

- To maximize the test takers' performances so that they come as close to their real competence as possible
- To vary tasks in an effort to find those tasks they feel most comfortable with
- To ensure that the content of the test (issues/subjects dealt with) do not penalize test takers who are not intimate with those subjects/issues
- To focus on the language production of the test takers- not to be influenced or swayed by their personalities
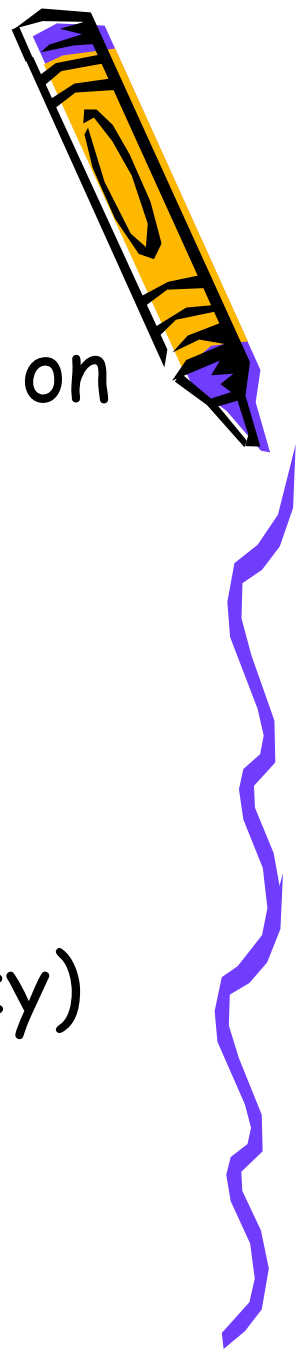
Validity and Reliability

- Validity: Tests assess what they are designed to assess (e.g. oral language skills and NOT subject area knowledge)

- Reliability
  - Inter-rater: Language production (data) is scored the same by DIFFERENT raters
  - Intra-rater: Language production (data) is scored the same by the SAME rater but at different times

Oral Test Task Types

- Narrative (e.g. story telling based on pictures, graph explanations, etc.)
- Interactive (e.g. Proficiency Interviews)
- Information Gap (e.g. describing a picture to another who draws it)
- Role Play (e.g. at the Travel Agency)
- Oral Translation-Interpretation

Curriculum vs Standards–Referenced Programs

- Curriculum-based programs…
  - consist of a series of related items to "cover"- often dictated by a textbook-curriculum
  - can be grammatical, functional, thematic, task-based, procedural, situational, etc. or a combination of the above

- Standards-referenced programs…
  - revolve around a set of articulated "behavioral objectives" which the course is designed to help learners achieve
  - are often institution-specific and reflect the institution's stated goals and assessed learner needs

Curriculum vs Standards–Referenced Exams

- Curriculum-based exams are designed to assess the learners' mastery of the content of the curriculum. The nature of the exam should reflect the content of the curriculum

- Standards-referenced exams are designed to assess the learners' mastery of the behavioral objectives set out for the course. These objectives are based on the institution's standards for particular levels/skills

Oral standards generally take the following into consideration
- nature of vocabulary used (range)
- pronunciation and comprehensibility
- syntactic complexity and accuracy
- fluency
- types of situations handled
- coherence
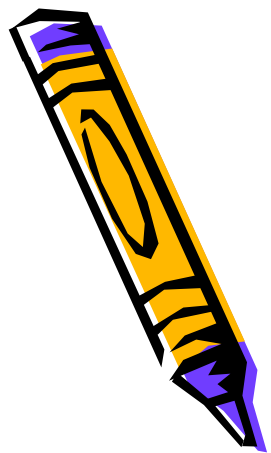
# Aural standards take into consideration…

- subjects/topics (from basic survival to abstract professional)
- source of input (e.g: face-face foreigner talk, mass media, academic lectures)
- length of input (e.g. isolated words, sentences, extended discourse)
- syntactic complexity
- vocabulary usage

Norming

- Objective- to ensure that at least 90% of all raters give same score to a given interview (inter-rater reliability)
- Procedure- play audiotapes of interviews where interlocutors are unknown to teachers
- Review ratings, discuss and continue- fine tuning of criteria descriptions for each level can occur during norming sessions
- Exam checks carried out by roaming monitors and spot audiotaped interview inspections
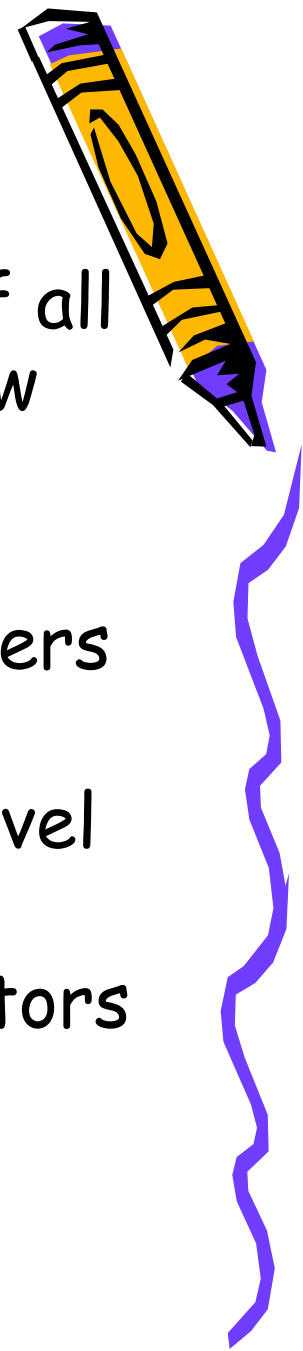
Interview training for oral assessment
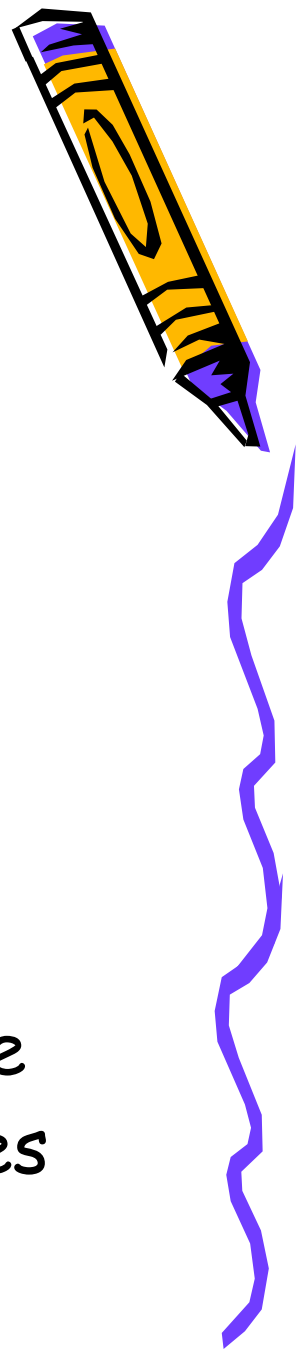
- Experienced interviewer demonstrates phases of interview
- Teachers put into groups
- Learners from different levels sent to each group
- 15-20 minute interviews conducted (audiotaped), followed by 10-15 minute small group discussion of interview style
- Full group meeting to discuss difficulties

Interview Phases

- Warm Up- put students at ease
- Level Check- determines the level the learner can comfortably function in - establishes base line
- Probes- requires learner to try to function at a level beyond the established baseline- if successful, a new, higher baseline is established. Probes continue till a "ceiling" (point when language breaks down) is established
- Wind Down- return to baseline level and end interview

Sample Interview Questions
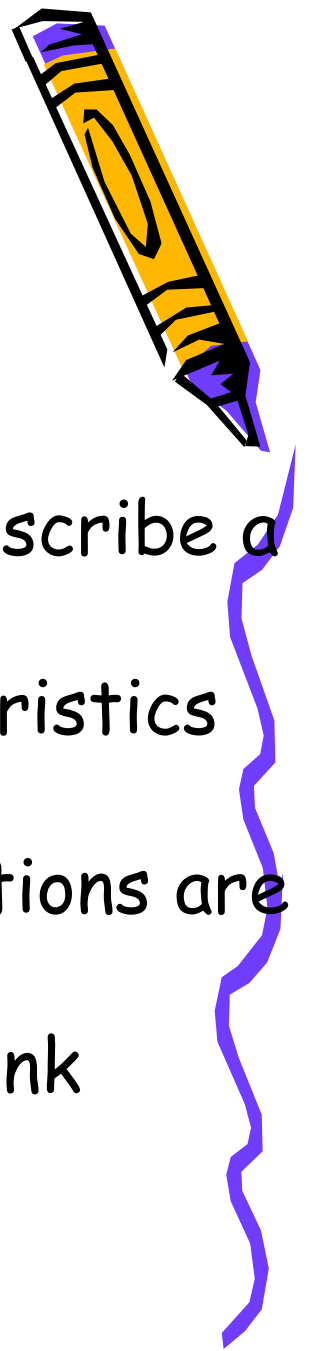
- Life-Personal (where do you live?)
- Life-Personal Concrete-Descriptive (Describe a wonderful place you have visited)
- Life Personal Abstract (What characteristics do you look for in a friend?)
- Non-Personal Concrete (What qualifications are necessary to practice medicine?)
- Non-Personal Abstract (Why do you think socialism failed in Eastern Europe?)

USEFUL REFERENCES

- Common European Framework of Reference for Languages: http://www.coe.int/T/E/Cultural_Cooperation/education/languages/Language_Policy

- Fulcher, G. (2003). <u>Testing Second Language Speaking</u>. London: Longman/Pearson Educational

- Davidson, F. and Lynch, B. (2002). <u>Testcraft: A Teacher's Guide to Writing and Using Language Test Specifications</u>. New Haven: Yale University Press.